Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2019) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

Postproceedings of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2018 (Ninth Annual Meeting of the BICA Society)

Virtual pet powered by a socially-emotional BICA

Anastasia A. Bogatyreva[a], Andrei D. Sovkov[a], Svetlana A. Tikhomirova[a], Anna R. Vinogradova[a] и Alexey V. Samsonovich[a,b]*

[a]*National Research Nuclear University "MEPhI" (Moscow Engineering Physics Institute) Kashirskoye, 31, Moscow, 115409, Russian Federation*
[b]*Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA 22030, USA*

Abstract

Cognitive architectures are used to build intelligent agents, and special attention in this area is drawn to emotion modelling. The purpose of this study is to compare the models, one of which is based on the machine learning algorithm, and the other on the socio-emotional cognitive architecture of eBICA. It is assumed that the BICA model will be much more effective in causing greater user empathy. The object of the model in this article is a pet – a penguin. For this purpose, a semantic map of pet's states was constructed by interviewing respondents and analyzing the collected data. Two methods were compared: reinforcement learning and BICA. The ratio of the results obtained using various methods was considered significant. The model based on eBICA has shown a better result in comparison to the one based on reinforcement learning. This article describes the strengths and the weaknesses of both methods. The comparison gives an idea of the effectiveness of the method based on the cognitive architecture of BICA, and will be useful for building intelligent agents.

Keywords: virtual agent; semantic space; emotional intelligence; reinforcement learning; cognitive architecture

* Corresponding authors.

E-mail addresses: A2Bog@list.ru, adsovkov@gmail.com, tikhomirova3112@yandex.ru, vinogradova.a.rom@gmail.com, asamsono@gmu.edu

1. Introduction

Currently, software modelling virtual environment (e.g., virtual and augmented reality systems) are becoming more common. For the convenience of human perception of the environment, in this environment it is advisable to place virtual assistants interacting with the user like living beings in the real world. Thus, developers are given the task of creating such virtual agents whose behavior was as plausible as possible from the point of view of a person who communicates with a pet. In this area, many studies and developments have been conducted and is being conducted; various architectures are being created that use a smartphone as an interface [1], various games are designed not only to entertain the customer (games, robots-toys [2-4]), but also to assist them (various intellectual advisers, expert agents, as well as robot assistants [5-7]). This research paper offers 2 methods: the first related to reinforcement learning and the second related to building a model based on a socio-emotional cognitive architecture [8]. These methods were chosen to test the main hypothesis of the article: agent-based behaviour with reinforcement training looks less natural than behaviour reproduced by an eBICA-based model. Thus, when comparing the behaviour of the two models based on the chosen methods, it is expected that the model based on eBICA will behave more realistically and, as a result, will be more popular among respondents.

2. Materials and methods

2.1. Description of the paradigm

The main goal of this paper is to implement two mathematical models of the behaviour of the actor and their research on human empathy with the help of experiments and statistical analysis. To achieve this goal, it was decided to implement a virtual pet - a penguin, which is surrounded by the following elements: a cave, a pool, a bowl with a picture of fish, a tower, several balls, and also a hand, which is the user's avatar. Screen with the penguin in the interior is presented. The animal is able to move around the area and interact with objects and hand.

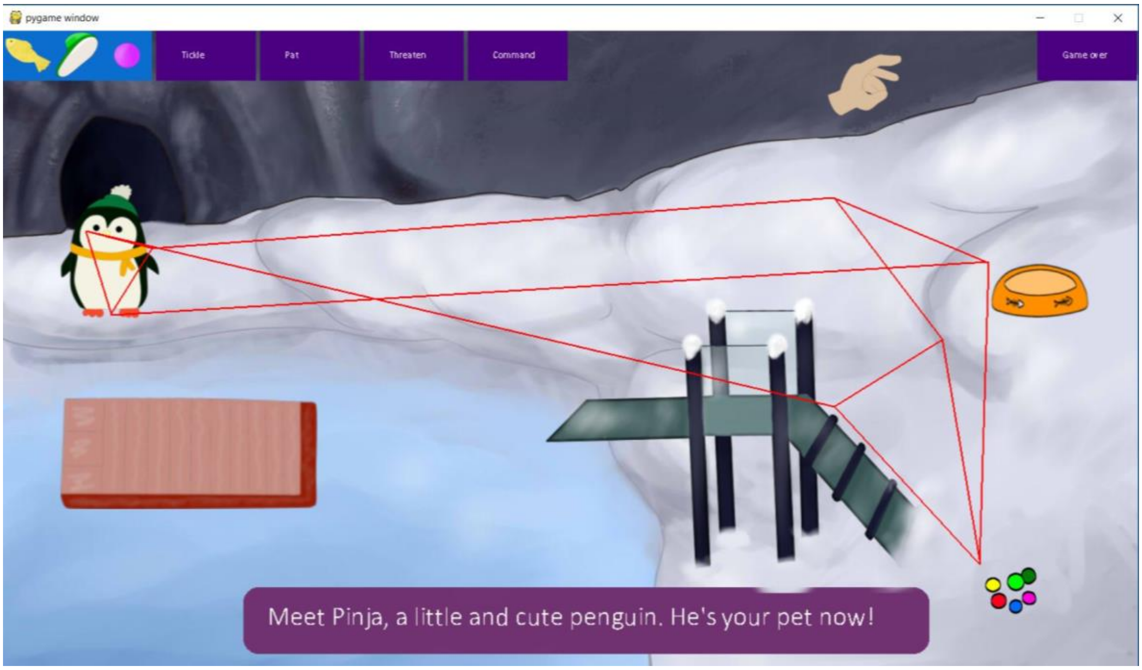

Fig. 1. Game window: a penguin surrounded. The red line indicates the graph in which the penguin can move. Leaf vertices are objects of the environment with which you can interact, the rest of the vertices are intermediate states as you move towards the object.

For example, it can swim in the pool, sleep in the cave, eat from the bowl, etc. The variety of hand's actions is also quite diverse: it can stroke, beat the animal, feed it and fill the bowl with food, as well as throw the ball. The movement occurs along the graph, where the vertices are the objects of the environment and points at the crossroads, and the edges connect these points (see Fig. 1).

It is important to note that the penguin must express various emotional states with its behavior (examples created by the authors are presented on Fig. 2).

Fig. 2. Examples of penguin's emotional states

2.2. Reinforcement learning

In order to start learning with reinforcements, it is necessary to define parameters, such as: agent, actions, rewards and target - and learning algorithm.

The animal itself is considered to be the agent; actions are its movements along the edges of the graph and attempts to interact with the environmental objects. The state of the penguin is represented by a vector, which consists of several elements: the values of the scales of satiety, vigour and fun, as well as the degree of the affinity with the person (the hand). An attempt to interact with the object is made when the distance between the animal and the object is very small. With each successful interaction, the penguin receives a reward. It is worth noting that the value of the reward can be both positive and negative, depending on the action performed. It also takes into account the fact that the hand, representing the user in the developed environment, can move and be the initiator of the contact with the animal, performing various actions: from negative to positive. The goal is to keep the agent's well-being scales farther from zero as long as possible, thus prolonging the game time.

For any finite Markov decision process, Q learning finds a policy that is optimal in the sense that it maximizes the expected value of the total reward over all successive steps, starting from the current state Q Learning can identify an optimal action-selection policy for any given decision process. "Q" names the function that returns the reward used to provide the reinforcement and can be said to stand for the "quality" of an action taken in a given state.

Q Learning, one of the methods of reinforcement learning used in machine learning, was used for training. The goal is a learning policy that explains why an agent chooses a particular action in these circumstances. If you're looking for what you need to do, it will help you given decision process. It is taken for a state.

The DQN algorithm was chosen as the algorithm, because it connects Q Learning with deep neural networks.

Deep Q network (DQN) [9] is represented by a multilayer neural network with three convolutional and two fully connected layers. At the output of this network for the current state st, the vector of values corresponding to the actions from the set is obtained.

The DQN agent interacts with the environment and updates the policy based on the actions taken (this is known as the on-policy learning algorithm). The Q-value for the state action is updated with an error corrected by learning speed α. Q-values are a possible reward received at the next time stage for taking measures at the st state, plus a γ discount on a future reward received from the next state observation. The value of Q-value itself occurs according to the formula (1). At time t + 1, all values are updated, especially the reward changes to rt + 1, and the current estimate to.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (1)$$

где:
- s_t и s_{t+1} - current and next state of the animal, respectively;
- a_t - the current action of the animal;
- r_{t+1} - reward;
- α, γ - positive coefficients.

The learning process is based on the experience replay technique [10], which stores experience, including state transitions, rewards and actions that are necessary data for Q Learning, and mini-batch files for updating neural networks.

2.3. Model based on eBICA

The basic model of pet behaviour in this paper is based on the construction of the eBICA (Biologically Inspired Cognitive Architecture), in which the emotional state of the agent is described using a semantic map. This model implies that the agent has an emotional component in addition to the scales of well-being (“sadness”, “anger”, “joy”, etc.).

In previous works [8], the states of the Aactor actors and the state shift when performing the action aaction were introduced. The semantic map was presented in the form of a complex space, where the valence value corresponded to the real axis, while the dominance map corresponded to the complex axis. In making the transition between the states of actors, a real parameter r was introduced, the role of which was to correct the shift of the emotional state during the interaction of actors.

In this work, the states of the actors (the penguin Atp and the human Ath) and the results of the interactions (the penguin aip and the human aih) are introduced as two-dimensional vectors of the semantic space. The voluntary interaction of the pet with the environment affects only the state of the penguin, the interaction of the person with the penguin changes both the assessment of the emotional state of the person and the penguin. These changes are described by formulas (2) - (3).

$$A^{t+1}{}_p = (1-r)A^t{}_p + r\, a^i{}_p \quad (2)$$

$$A^{t+1}{}_h = (1-r)A^t{}_h + r\, a^i{}_h \quad (3)$$

The penguin’s preferences in the choice of actions when interacting with the environment are determined as follows: in a given state for the i-th possible interaction, aip and ayh are estimated using the above calculation formulas. data states. For falling distances, the average value of d * is calculated, and as possible interactions, ai are chosen such that. Further action selection is performed randomly.

The axes of the semantic space were determined by interviewing the respondents and analyzing the principal components (principal components analysis). The numerical values of the interaction vectors were determined using a survey - a more detailed description of the process and the results are presented in part 3, experiment No.

3. **Results and analysis**

3.1. Experiment №1

To build a semantic map, an experiment was conducted in which 25 people participated (16 men and 9 women), most of whom were students. Each of them was asked to rate 13 animated penguin images by the following criteria: interest, serenity, trust, anxiety, agitation, anger, resentment, aggression, humility, pleasure, love, disappointment and compassion, every of which was rated on a scale from 0 to 10, where 0 is the absence of emotion in animation, and 10 is the maximum expression of emotion. For the collected data, the analysis of the main components was carried out. As a result, the following two-dimensional semantic map with valence and dominance axes was obtained (see Fig. 3).

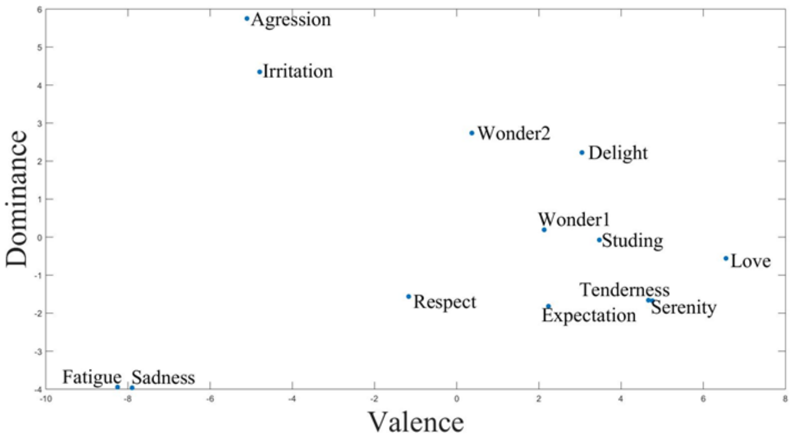

Fig.3. Semantic map

The results of this experiment confirmed the hypothesis that when building a two-dimensional semantic map, it is preferable to use the valence and dominance axes, since they provide the greatest variation in the emotional states of the pet.

3.2. Experiment № 2

To facilitate the construction of a mathematical model, an experiment was conducted in which 12 students took part (10 men and 2 women). Each of the subjects was asked to evaluate different situations on two scales of valence & dominance ranging from -5 to +5. +5 valence meant a positive reaction, and -5 valence - a negative reaction. +5 dominance corresponded to dominance, and -5 submission. The survey was divided into two parts: in the first section, the probationers were asked to pretend to be the pet and evaluate the various actions of the owner in relation to him, and in the second, the probationers pretended to be the owner and evaluated the reactions of the penguin. In this survey, the subjects were shown 27 fragments from the game. The results of the analysis of the data obtained are shown in Fig. 4.

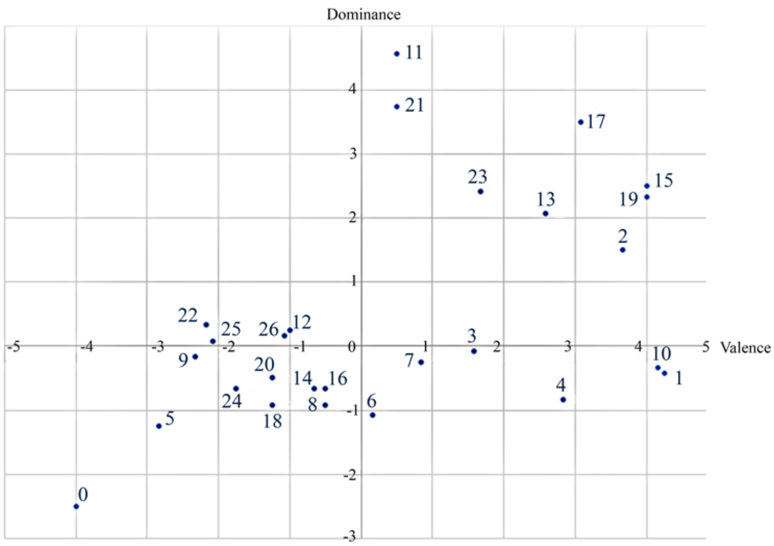

Fig 4. The position of the points of the emotional state of the penguin relative to the neutral state on the semantic axis.

The points indicated on the semantic map correspond to the situations presented below.

• The owner hits you with a slipper.
• The owner feeds you with fish from the hands.
• The owner let you play with the balls.
• The owner tickles you.
• The owner pets you on the head.
• The owner scolds you for bad behavior.
• The owner orders you to ride a hill.
• The owner orders you to swim on the mattress in the pool.
• The owner sends you to sleep.
• The owner forcibly wakes you up.
• The owner sends you to eat from the bowl.
• You hit a penguin with a slipper. He dutifully accepted the punishment.
• You hit a penguin with a slipper. He aggressively responded to this action.
• You offer a penguin to eat fish from your hands. He eats it with pleasure.
• You offer a penguin to eat fish from your hands. He continues to go about his business / runs away.
• You offer a penguin to play with balls. He eagerly starts playing.
• You offer a penguin to play with balls. He continues to go about his business / runs away.
• You begin to tickle the penguin. He rejoices caress.
• You begin to tickle the penguin. He snaps at you / runs away.
• You start petting the penguin on the head. He rejoices caress.
• You start petting the penguin on the head. He snaps at you / runs away.
• You chastise a penguin for bad behavior. He dutifully accepts the punishment.
• You chastise a penguin for bad behavior. He snaps at you.
• You order the penguin to go down the slide / swim on the mattress / go to sleep. He obeys.
• You order the penguin to go down the slide / swim on the mattress / go to sleep. He continues to go about his business.
• You are trying to wake the penguin. He does not respond to you and continues to sleep.
• You can not find a penguin in the aviary, as he hid in his cave.

Fig. 5. Fragment of the game - “Interaction of a penguin with a slide”.

Thus, conducting these experiments allowed us to construct vectors used to construct a mathematical model.

3.3. Experiment №3

The experiment involved 13 people. Participants were asked to interact with two pet models, one of which was based on pure reinforcement training, the other on a model using cognitive architecture. Each of the subjects was asked to interact with the penguin as they would interact with their own pet, the subjects did not know about the features of each of the models. All user interactions, as well as the interactions of the penguin with the environment, were recorded in a log file.

After the experiment, subjects underwent a survey in which, on a scale from 1 to 10, the impressions of both models were evaluated, as well as their realism.

3.4. Results

The following average values were obtained when processing the survey data (table 1):

Table 1. Results

	BICA	RL
Rate impressions about the model, 1 - did not like it, 10 - really liked it	7.92	6.31
How realistic was this model? 1 - unrealistic, 10 - realistic	7.15	6.08

Dispersion analysis of the results showed the value of p-value = 0.0436 for the first survey and p-value = 0.1613 for the second survey. The high value of p-value in the second survey, and as a result, a high probability of acceptance may be due to the fact that the subjects could not bring the penguin far from the average state, noticing the clearer manifestations of the architecture used. The model at BICA also suggested that when the valence is low, the penguin stops executing user commands, and the user receives messages about the penguin’s failure to fully interact with it, however, users did not reach the low values of the penguin valence during the experiment. In general, however, it is possible to speak of a statistically significant superiority of the model with BICA over the conventional model in training with reinforcement in realism and a challenge to empathy among users.

4. Discussion

In this paper, a prototype of an emotional pet penguin in the environment was implemented. With the help of a respondents poll and analysis of the main components, a semantic map was constructed. Comparison of training models with reinforcements and at BICA have been described. However, in this paper one of the most simple levels of the user's emotional interaction with the pet was considered. In the future, the pet must respond differently to the actions of a person depending on the relationship between them. This will make it more realistic in the eyes of the user who will interact with it. In addition, the penguin will show more complex emotions such as resentment, disappointment, remorse, awe and others.

5. Acknowledgments

This work was supported by the Russian Science Foundation, Grant No 18-11-00336. The authors are grateful to the students of the NRNU MEPhI for lively discussions and participation in experiments.

References

[1] Antonio Chella, Rosario Sorbello, Giovanni Pilato, Giorgio Vassallo, and Marcello Giardina (2013) "A New Humanoid Architecture for Social Interaction between Human and a Robot Expressing Human-Like Emotions Using an Android Mobile Device as Interface" *Biologically Inspired Cognitive Architectures 2012. Advances in Intelligent Systems and Computing, vol 196*: 95–103.

[2] Daniil A. Azarnov, Artur A. Chubarov and Alexei V. Samsonovich. (2017) "Virtual Actor with Social-Emotional Intelligence" *Procedia Computer Science 123 (2018)*: 76–85.

[3] Agnese Augello, Giovanni Pilato and Salvatore Gaglio. (2009) "A conversational agent to support decisions in SimCity like games" *Proceedings – IEEE ICSC* (2009): 367–372.

[4] Jérôme Monceaux, Joffrey Becker, Céline Boudier, Alexandre Mazel (2009). "Demonstration – First Steps in Emotional Expression of the Humanoid Robot Nao". *ICMI-MLMI'09*, November 2–4, 2009: 235-236.

[5] Antonio Chella, Giovanni Pilato, Rosario Sorbello, Giorgio Vassallo, Francesco Cinquegrani, and Salvatore Maria Anzalone (2008). "An Emphatic Humanoid Robot with Emotional Latent Semantic Behavior" *SIMPAR 2008*: 234-245

[6] Giovanni Pilato, Giorgio Vassallo, Agnese Augello, Maria Vasile, and Salvatore Gaglio (2005). "Expert chat-bots for cultural heritage"*Intelligenza Artificiale Journal* vol.2 (2005): 25-31

[7] Agnese Augello, Giovanni Pilato, and Salvatore Gaglio (2010). "Intelligent Advisor Agents in Distributed Environments" *Information Retrieval and Mining in Distributed Environments*: 109-124.

[8] Samsonovich, A. V. (2013). "Emotional biologically inspired cognitive architecture" *Biologically Inspired Cognitive Architecture* (2013) **6**: 109-125

[9] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Demis Hassabis (2015). "Human-level control through deep reinforcement learning" *Nature* volume 518 (26 February 2015): 529–533

[10] Long-Ji Lin (2013). "Reinforcement Learning for Robots Using Neural Networks". *Doctoral Dissertation Reinforcement learning for robots using neural networks Carnegie Mellon University Pittsburgh, PA, USA ©1992*